Upper bound by Kolmogorov complexity for the probability in computable POVM measurement

Kohtaro Tadaki

IMAI Quantum Computation and Information Project, ERATO,
Japan Science and Technology Corporation,
Daini Hongo White Bldg. 201, 5-28-3, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
TEL: +81-3-3818-3314 FAX: +81-3-3818-3285
E-mail: tadaki@qci.jst.go.jp

Abstract. We apply algorithmic information theory to quantum mechanics in order to shed light on an algorithmic structure which inheres in quantum mechanics.

There are two equivalent ways to define the (classical) Kolmogorov complexity $K(s)$ of a given classical finite binary string s. In the standard way, $K(s)$ is defined as the length of the shortest input string for the universal self-delimiting Turing machine to output s. In the other way, we first introduce the so-called universal probability m, and then define $K(s)$ as $-\log_2 m(s)$ without using the concept of program-size. We generalize the universal probability to a matrix-valued function, and identify this function with a POVM (positive operator-valued measure). On the basis of this identification, we study a computable POVM measurement with countable measurement outcomes performed upon a finite dimensional quantum system. We show that, up to a multiplicative constant, $2^{-K(s)}$ is the upper bound for the probability of each measurement outcome s in such a POVM measurement. In what follows, the upper bound $2^{-K(s)}$ is shown to be optimal in a certain sense.

Key words: algorithmic information theory, universal probability, POVM, computability, quantum Kolmogorov complexity

1 Introduction

Algorithmic information theory is a theory of program-size complexity which has precisely the formal properties of classical information theory. In algorithmic information theory, the *program-size complexity* (or *Kolmogorov complexity*) $K(s)$ of a finite binary string s is defined as the length of the shortest binary input for the universal self-delimiting Turing machine to output s. The concept of program-size complexity plays an important role in characterizing the randomness of a finite or infinite binary string. In this paper we extend algorithmic information theory to quantum region in order to throw light upon an algorithmic feature of quantum mechanics. We show that Kolmogorov complexity gives the upper bound for the probability of each measurement outcome in a computable POVM measurement with countable outcomes performed upon a finite dimensional quantum system.

1.1 Main result

In this paper, we consider a quantum measurement performed upon a *finite dimensional* quantum system. A *positive operator-valued measure* (POVM) is a collection $\{E(m)\}$ of positive semi-definite Hermitian matrices which satisfies $\sum_m E(m) = I$ where I is the identity matrix. Each $E(m)$ is called a *POVM element* of this POVM. In general, the statistics of outcomes in a quantum measurement are described by a POVM $\{E(m)\}$. The label m refers to the measurement outcomes that may occur in the experiment. If the state of the quantum system is described by a normalized vector $|\psi\rangle$ immediately before the measurement, then the probability that result m occurs is given by $\langle\psi|E(m)|\psi\rangle$. On the other hand, if the ensemble of the states of the quantum system is described by a density matrix ρ immediately before the measurement, then the probability that result m occurs is given by $\mathrm{tr}(\rho E(m))$. A POVM measurement is a generalization of a *projective measurement* which is described by an observable. The number of outcomes in a POVM measurement can be more than the dimension of the state space of the quantum system being measured, whereas the number of outcomes in a projective measurement cannot. In this paper, we relate an argument s of $K(s)$ to an outcome which may occur in the quantum measurement performed upon a finite dimensional quantum system. Since $K(s)$ is defined for all finite binary strings s, the countable outcomes have to be available in the corresponding quantum measurement. Thus we deal with a POVM measurement and not a projective measurement. (See e.g. [11, 12] for the details of POVM measurements.)

We say a POVM is *computable* if one can compute all its POVM elements to any desired degree of precision, and a POVM measurement is said to be *computable* if it is described by a computable POVM. Our main result is as follows: Let $\{R(s)\}$ be a computable POVM on a finite dimensional quantum system whose each element is labeled by a finite binary string. Then there exists an integer d such that, for all density matrix ρ and all finite binary string s,

$$K(s) - d \leq -\log_2 \mathrm{tr}(\rho R(s)), \tag{1}$$

and also there exists a real number $c > 0$ such that, for all density matrix ρ and all finite binary string s,

$$\mathrm{tr}(\rho R(s)) \leq c\, P(s). \tag{2}$$

Here $P(s)$ is the probability that the (classical) universal self-delimiting Turing machine halts and outputs s when it starts on the program tape filled with an infinite binary string generated by infinitely repeated tosses of a fair coin.

The inequality (1) states that, up to an additive constant, $K(s)$ is the lower bound for the $-\log_2$ of the probability of each measurement outcome s in a computable POVM measurement with countable outcomes performed upon a finite dimensional quantum system, i.e., $2^{-K(s)}$ is the upper bound for the probability of each outcome s up to a multiplicative constant. On the other hand, the inequality (2) states that, up to a multiplicative constant, $P(s)$ is the upper bound for the probability of each measurement outcome s in the same measurement. Note that the inequalities (1) and (2) are equivalent to each other.

The computability of a POVM measurement is thought to be intrinsic in the case where one performs the measurement in order to extract a valuable information from a quantum system because in such a case one has to be able to compute to any desired degree of precision all POVM elements of the POVM which describes the measurement. Hence, when one wants to extract a valuable information from a finite dimensional quantum system through a POVM measurement with countable outcomes, one faces with the limitation given by the inequality (1) (equivalently by (2)).

Especially, the inequality (2) is interesting. Since $P(s)$ is a probability which results from infinitely repeated tosses of a fair coin, $P(s)$ is just a classical probability. In the case where ρ

is a pure state, the inequality (2) states that a purely quantum mechanical probability bounded from above by a purely classical probability up to a multiplicative constant when one performs a computable POVM measurement with countable outcomes upon a finite dimensional quantum system in the pure state ρ.

The inequalitis (1) and (2) are obtained through a generalization of the so-called *universal probability* to a matrix-valued function. The Kolmogorov complexity $K(s)$ of a finite binary string s is originally defined using the concept of program-size. However, there is another way to define $K(s)$ without referring to such a concept, that is, we first introduce a universal probability m, and then define $K(s)$ as $-\log_2 m(s)$. The universal probability is a function from the set of finite binary strings to the open interval $(0, 1)$. In this paper we generalize the universal probability to a matrix-valued function while keeping the domain of definition the set of finite binary strings. Then this generalized universal probability is identified with an analogue of a POVM, and is called a *universal semi-POVM*. The inequalities (1) and (2) naturally follow from this identification.

1.2 Related works

Our aim is to generalize algorithmic information theory in order to understand the algorithmic feature of quantum mechanics. There are related works whose purpose is mainly to define the information content of an individual pure quantum state, i.e., to define the *quantum Kolmogorov complexity* of the quantum state [13, 2, 8], while we will not make such an attempt in this paper.

As we mentioned above, $K(s)$ can be defined as the $-\log_2$ of the universal probability without using the concept of program-size. [8] took this approach in order to define the information content of a pure quantum state. [8] first generalized the universal probability to a matrix-valued function $\boldsymbol{\mu}$, called *quantum universal semi-density matrix*. The $\boldsymbol{\mu}$ is a function which maps any positive integer N to an $N \times N$ positive semi-definite Hermitian matrix $\boldsymbol{\mu}(N)$ with its trace less than or equal to one. [8] proposed to regard $\boldsymbol{\mu}(N)$ as an analogue of a density matrix of a quantum system called *semi-density matrix*. Then, in order to measure the information content of a pure quantum state $|\psi\rangle \in \mathbb{C}^N$, [8] introduced the *quantum algorithmic entropies* $\underline{H}(|\psi\rangle)$ and $\overline{H}(|\psi\rangle)$ as $-\log_2\langle\psi|\boldsymbol{\mu}(N)|\psi\rangle$ and $-\langle\psi|(\log_2 \boldsymbol{\mu}(N))|\psi\rangle$, respectively. In general, the trace of a density matrix has to be equal to one. If the trace of $\boldsymbol{\mu}(N)$ is equal to one, then the quantity $\langle\psi|\boldsymbol{\mu}(N)|\psi\rangle$ in the definition of $\underline{H}(|\psi\rangle)$ has the meaning of the probability that the outcome is 'yes' when one performs the projective measurement described by the projector $|\psi\rangle\langle\psi|$ upon the quantum system in the mixed state $\boldsymbol{\mu}(N)$. However, the trace of $\boldsymbol{\mu}(N)$ is not equal to one for all but finitely many N because of its universality. (This fact is implicitly mentioned in [8]. For completeness, we include a proof of this fact in Appendix A, in addition to the definition of $\boldsymbol{\mu}$.)

In quantum mechanics, what is represented by a matrix is either a quantum state or a measurement operator. In this paper we generalize the universal probability to a matrix-valued function in different way from [8], and identify it with an analogue of a POVM. We do not stick to defining the information content of a quantum state. Instead, we focus our thoughts on applying algorithmic information theory to quantum mechanics in order to shed light on an algorithmic structure of quantum mechanics. In this line we have the above inequalities (1) and (2).

In each of [13] and [2], the quantum Kolmogorov complexity of a qubit string was defined as a quantum generalization of the standard definition of classical Kolmogorov complexity; the length of the shortest input for the universal decoding algorithm U to output a finite binary string. Both [13] and [2] adopt the *universal quantum Turing machine* as a universal decoding algorithm U to output a quantum state in their definition. However, there is a difference between [13] and [2] with respect to the object which is allowed as an input to U. That is, [13] can only allow a

classical binary string as an input, whereas [2] can allow any qubit string. The works [13], [2], and [8] are closely related to one another as shown in each of these works. In comparison with our work, since our work is, in essence, based on a generalization of the universal probability, the work [8] is more related to our work than the works [13] and [2]. These two works may be related to our work via the work [8].

1.3 Organization of the paper

We begin in Section 2 with some basic definitions, and review some results of algorithmic information theory. In Section 3, we prove the inequalities (1) and (2) via the introduction of a universal semi-POVM as a generalization of the universal probability. In Section 4, we consider the optimality of our upper bound $2^{-K(s)}$ and $P(s)$ for the probability of each measurement outcome s. Finally, we study some other properties of a universal semi-POVM in Section 5.

2 Preliminaries

2.1 Notation

We start with some notation about numbers and matrices which will be used in this paper.

$\mathbb{N} \equiv \{0, 1, 2, 3, \dots\}$ is the set of natural numbers, and $\mathbb{N}^+$ is the set of positive integers. $\mathbb{Z}$ is the set of integers. $\mathbb{Q}$ is the set of rational numbers, and $\mathbb{Q}^+$ is the set of positive rational numbers. $\mathbb{R}$ is the set of real numbers, and $\mathbb{C}$ is the set of complex numbers. $\mathbb{C}_Q$ is the set of the complex numbers in the form of $a + ib$ with $a, b \in \mathbb{Q}$. We define $-\log_2 0$ as ∞.

We fix N to be any one positive integer throughout this paper. For each matrix A, A^T is the transpose of A and $A^\dagger$ is the adjoint of A. For each $K \subset \mathbb{C}$, $M_N(K)$ is the set of the $N \times N$ matrices whose elements are in K, and K^N is the set of column vectors consist N complex numbers in K. For each $x = (x_1, x_2, \dots, x_N)^T \in \mathbb{C}^N$, $\|x\|$ is defined as $(|x_1|^2 + |x_2|^2 + \cdots + |x_N|^2)^{1/2}$. For each $A, B \in M_N(\mathbb{C})$, $[A, B]$ is defined as $AB - BA$. For each $A \in M_N(\mathbb{C})$, $\|A\|$ is the *operator norm* of A, and $\operatorname{tr} A$ denotes the *trace* of A. The *identity matrix* in $M_N(\mathbb{C})$ is denoted by I. $U(N)$ is the set of $N \times N$ unitary matrices. $\mathrm{Her}(N)$ is the set of $N \times N$ Hermitian matrices. For each $A, B \in \mathrm{Her}(N)$, we write $A \leqslant B$ if $B - A$ is positive semi-definite, and write $A < B$ if $B - A$ is positive definite. Note that the relation $\leqslant$ on $\mathrm{Her}(N)$ is a partial order. In this paper we will frequently use the property: $\|A\| \leq \varepsilon \iff -\varepsilon I \leqslant A \leqslant \varepsilon I$ for any $\varepsilon \geq 0$ and any $A \in \mathrm{Her}(N)$. We say ρ is a *density matrix* if $0 \leqslant \rho \in \mathrm{Her}(N)$ and $\operatorname{tr}(\rho) = 1$. $\mathrm{Her}_{\mathbb{Q}}(N)$ is the set of $N \times N$ Hermitian matrices whose elements are in $\mathbb{C}_Q$. $\operatorname{diag}(x_1, \dots, x_N)$ is the diagonal matrix whose (i, i)-elements is x_i.

Let S be any set, and let $f, g \colon S \to \mathrm{Her}(N)$. Then we write $f(x) = g(x) + O(1)$ if there is a real number $c > 0$ such that, for all $x \in S$, $\|f(x) - g(x)\| \leq c$. We also write $f(x) \sim g(x)$ if there is a real number $c > 0$ such that, for all $x \in S$, $c\, f(x) \leqslant g(x)$ and $c\, g(x) \leqslant f(x)$.

$\Sigma^* \equiv \{\lambda, 0, 1, 00, 01, 10, 11, 000, 001, 010, \dots\}$ is the set of finite binary strings where λ denotes the *empty string*, and Σ^* is ordered as indicated. We identify any string in Σ^* with a natural number in this order, that is, we consider $\varphi \colon \Sigma^* \to \mathbb{N}$ such that $\varphi(s) = 1s - 1$ where the concatenation $1s$ of strings 1 and s is regarded as a dyadic integer, and then we identify s with $\varphi(s)$. For any $s \in \Sigma^*$, $|s|$ is the *length* of s. A subset S of Σ^* is called a *prefix-free set* if no string in S is a prefix of another string in S.

For each $F \colon \Sigma^* \to M_N(\mathbb{C})$, we say F is *computable* if there exists a total recursive function $G \colon \Sigma^* \times \mathbb{N} \to M_N(\mathbb{C}_Q)$ such that, for all $s \in \Sigma^*$ and all $k \in \mathbb{N}$, $\|F(s) - G(s, k)\| < 2^{-k}$.

2.2 Algorithmic information theory

In the following we review some definitions and results of algorithmic information theory [6, 7]. We assume that the reader is familiar with algorithmic information theory in addition to computability theory.

A *computer* is a partial recursive function $C\colon \Sigma^* \to \Sigma^*$ whose domain of definition is a prefix-free set. For each computer C and each $s \in \Sigma^*$, $K_C(s)$ is defined as $\min\left\{\, |p| \;\middle|\; p \in \Sigma^* \,\&\, C(p) = s \,\right\}$. A computer U is said to be *optimal* if for each computer C there exists a constant $\mathrm{sim}(C)$ with the following property: if $C(p)$ is defined, then there is a p' for which $U(p') = C(p)$ and $|p'| \leq |p| + \mathrm{sim}(C)$. It is then shown that there exists a computer which is optimal. We choose any one optimal computer U as the standard one for use throughout the rest of this paper, and we define $K(s) \equiv K_U(s)$, which is referred to as the *information content* of s, the *program-size complexity* of s, or the *Kolmogorov complexity* of s. For each $s \in \Sigma^*$, $P(s)$ is defined by $P(s) \equiv \sum_{U(p)=s} 2^{-|p|}$. The class of computers is equal to the class of functions which are computed by *self-delimiting Turing machines*. A self-delimiting Turing machine has a program tape and a work tape. The program tape is infinite to the right, while the work tape is inifinite in both directions. The machine starts with an input string on its program tape and the work tape blank. When the machine halts, the output string is put on the work tape. (For the details of self-delimiting Turing machine, see [6].) A self-delimiting Turing machine is called *universal* if it computes an optimal computer. Let M_U be a universal self-delimiting Turing machine which computes U. Then $P(s)$ is the probability that M_U halts and outputs s when M_U starts on the program tape filled with an infinite binary string generated by infinitely repeated tosses of a fair coin.

A universal probability is defined through the following two definitions.

Definition 2.1. *For any $r\colon \Sigma^* \to [0,\infty)$, we say that r is a lower-computable semi-measure if r satisfies the following two conditions:*

(i) $\sum_{s\in\Sigma^*} r(s) \leq 1$.

(ii) There exists a total recursive function $f\colon \mathbb{N}\times\Sigma^* \to \mathbb{Q}$ such that, for each $s \in \Sigma^*$, $\lim_{n\to\infty} f(n,s) = r(s)$ and $\forall\, n \in \mathbb{N}\ \ f(n,s) \leq f(n+1,s)$.

Definition 2.2. *Let m be a lower-computable semi-measure. We say that m is a universal probability if for any lower-computable semi-measure r, there exists a real number $c > 0$ such that, for all $s \in \Sigma^*$, $c\,r(s) \leq m(s)$.*

Then the following theorem holds.

Theorem 2.3. *Both $2^{-K(s)}$ and $P(s)$ are universal probabilities.*

By Theorem 2.3, we see that, for any universal probability m,

$$K(s) = -\log_2 m(s) + O(1). \tag{3}$$

Especially we have $K(s) = -\log_2 P(s) + O(1)$. Any universal probability is not computable, which corresponds to the uncomputability of $K(s)$. Moreover we can show the following, from which the uncomputability of a universal probability follows.

Theorem 2.4. *Let m be a universal probability, and let $f\colon \mathbb{N}\to\mathbb{Q}^+$ and $\tau\colon \mathbb{N}\to\Sigma^*$. Suppose that both f and τ are total recursive functions, and $m(\tau(n)) \leq f(n)$ for all $n \in \mathbb{N}$. Then $\inf_{s\in\Sigma^*} f(n) > 0$.*

The information theoretic feature of algorithmic information theory can be developed as follows. We choose any one computable bijection $< s,t >$ from $(s,t) \in \Sigma^* \times \Sigma^*$ to Σ^*. Let $s,t \in \Sigma^*$. The *joint information content* $K(s,t)$ of s and t is defined as $K(s,t) \equiv K(< s,t >)$. We then define the *relative information content* $K(s|t)$ of s relative to t by the equation

$$K(s|t) \equiv K(t,s) - K(t).$$

Finally we define the *mutual information content* $K(s:t)$ of s and t by the equation

$$K(s:t) \equiv K(t) - K(t|s) \equiv K(s) + K(t) - K(s,t).$$

Then, without referring to the concept of program-size, [7] proved the following relations using the fact that $2^{-K(s)}$ is a universal probability.

Theorem 2.5.

(i) $K(s,t) = K(t,s) + O(1)$.

(ii) $K(s:t) = K(t:s) + O(1)$.

(iii) $K(s:s) = K(s) + O(1)$.

(iv) $\exists\, c \in \mathbb{R}\ \ \forall\, s,t \in \Sigma^*\ \ c \leq K(s|t)$.

(v) $\exists\, c \in \mathbb{R}\ \ \forall\, s,t \in \Sigma^*\ \ c \leq K(s:t)$.

(vi) $K(s:t) = K(t:s) + O(1)$.

(vii) $K(s:s) = K(s) + O(1)$.

(viii) $K(s:\lambda) = O(1)$.

Thus algorithmic information theory has the formal properties of classical information theory.

3 Generalization of universal probability to POVM

In this section we generalize a universal probability to a matrix-valued function. Based on this generalization, we prove our main result: Theorem 3.9.

Definition 3.1. *We say R is a semi-POVM on Σ^* if R is a mapping from Σ^* to $\mathrm{Her}(N)$ which satisfies $0 \leqslant R(s)$ for all $s \in \Sigma^*$ and $\sum_{s\in\Sigma^*} R(s) \leqslant I$. We say R is a POVM on Σ^* if R is semi-POVM on Σ^* and $\sum_{s\in\Sigma^*} R(s) = I$.*

Let Q be a POVM on Σ^*. The POVM measurement described by Q is performed upon a finite dimensional quantum system, and gives one of countable measurement outcomes, which are represented by finite binary strings.

Given R: semi-POVM on Σ^*, it is easy to convert R into a POVM on Σ^* by appending an appropriate positive semi-definite matrix to R. Let $\Omega = \sum_{s\in\Sigma^*} R(s)$, and then we define $Q\colon \Sigma^* \to \mathrm{Her}(N)$ by $Q(\lambda) = I - \Omega$ and $Q(s') = R(s)$ for each $s \in \Sigma^*$ where s' is the successor of s. Then Q is a POVM on Σ^*. Thus a semi-POVM on Σ^* has a physical meaning in the same way as a POVM on Σ^*.

Definition 3.2. *We say R is a lower-computable semi-POVM if R is a semi-POVM on Σ^* and there exists a total recursive function $f\colon \mathbb{N} \times \Sigma^* \to \mathrm{Her}_{\mathbb{Q}}(N)$ such that, for each $s \in \Sigma^*$, $\lim_{n\to\infty} f(n,s) = R(s)$ and $\forall\, n \in \mathbb{N}\ \ f(n,s) \leqslant R(s)$.*

In the case where $N = 1$, Definition 3.2 exactly results in the definition of a lower-computable semi-measure. For the handiness, we do not reqiure in the above definition that the $f(n, s)$ conversing to $R(s)$ is non-decreasing (i.e., $f(n, s) \leqslant f(n + 1, s)$). However, we can equivalently assume that the $f(n, s)$ is non-decreasing in the definition. See Appendix B for its proof.

The following is a key theorem for our main result.

Theorem 3.3. *If R is a lower-computable semi-POVM, then the mapping $\Sigma^* \ni s \longmapsto \frac{1}{N}\|R(s)\|$ is a lower-computable semi-measure.*

Proof. Let $r\colon \Sigma^* \to [0, \infty)$ with $r(s) = \frac{1}{N}\|R(s)\|$. Note that $\|A\| \leq \operatorname{tr} A$ for any positive semi-definite A. Thus, since $0 \leqslant R(s)$ for all $s \in \Sigma^*$ and $\sum_{s\in\Sigma^*} R(s) \leqslant I$, we see that $\sum_{s\in\Sigma^*} r(s) \leq \frac{1}{N} \operatorname{tr} \sum_{s\in\Sigma^*} R(s) \leq \frac{1}{N} \operatorname{tr} I = 1$. Thus the condition (i) in Definition 2.1 holds for r.

Next we show that the condition (ii) in Definition 2.1 holds for r. Since R is a lower-computable semi-POVM, there exists a total recursive function $f\colon \mathbb{N} \times \Sigma^* \to \mathrm{Her}_{\mathbb{Q}}(N)$ such that for each $s \in \Sigma^*$, $\lim_{n\to\infty} f(n, s) = R(s)$ and $\forall n \in \mathbb{N}\ \ f(n, s) \leqslant R(s)$. From the definition of the operator norm, $\|f(n, s)\|$ is the supremum of $\langle\psi|f(n, s)|\psi\rangle/\langle\psi|\psi\rangle$ such that $|\psi\rangle \neq 0$ and $|\psi\rangle \in \mathbb{C}^N$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, it is easy to see that $\|f(n, s)\|$ is equal to the supremum of $\langle\psi|f(n, s)|\psi\rangle/\langle\psi|\psi\rangle$ such that $|\psi\rangle \neq 0$ and each component of $|\psi\rangle$ is a complex number in the form of $a + ib$ with $a, b \in \mathbb{Z}$. Thus, given $n \in \mathbb{N}$ and $s \in \Sigma^*$, one can generate a sequence of rational numbers $p_1, p_2, \ldots$ such that $p_1 \leq p_2 \leq \cdots \leq \|f(n, s)\|$ and $\lim_{m\to\infty} p_m = \|f(n, s)\|$. On the other hand, using the property $A \leqslant B \Longrightarrow \|A\| \leq \|B\|$, we have $\|f(n, s)\| \leq \|R(s)\|$ and $\lim_{n\to\infty} \|f(n, s)\| = \|R(s)\|$. Hence, given $s \in \Sigma^*$, one can generate a sequence of rational numbers $x_1, x_2, \ldots$ such that $x_1 \leq x_2 \leq \cdots \leq \|R(s)\|$ and $\lim_{n\to\infty} x_n = \|R(s)\|$. Therefore the condition (ii) in Definition 2.1 holds for r. Hence r is a lower-computable semi-measure. □

Definition 3.4. *Let M be a lower-computable semi-POVM. We say that M is a universal semi-POVM if for each lower-computable semi-POVM R, there exists a real number $c > 0$ such that for all $s \in \Sigma^*$, $c\,R(s) \leqslant M(s)$.*

In the case where $N = 1$, Definition 3.4 exactly results in the definition of a universal probability. The use of the partial order $\leqslant$ for the purpose of generalizing lower-computable semi-measure and universal probability to matrix-valued functions is suggested in [8]. Note that if M is a universal semi-POVM then, for all $s \in \Sigma^*$, $M(s)$ is positive definite.

A universal semi-POVM may have a simple form as the following theorem says.

Theorem 3.5. *If m is a universal probability, then the mapping $\Sigma^* \ni s \longmapsto m(s)I$ is a universal semi-POVM.*

Proof. Let $M\colon \Sigma^* \to \mathrm{Her}(N)$ with $M(s) = m(s)I$. Since m is a lower-computable semi-measure, it is obvious that M is a lower-computable semi-POVM. Suppose that R is a lower-computable semi-POVM. By Theorem 3.3, the mapping $\Sigma^* \ni s \longmapsto \frac{1}{N}\|R(s)\|$ is a lower-computable semi-measure. Thus, since m is a universal probability, there is $c > 0$ such that, for all $s \in \Sigma^*$, $c\frac{1}{N}\|R(s)\| \leq m(s)$. Therefore we have $\frac{c}{N}R(s) \leqslant m(s)I$ for all $s \in \Sigma^*$. Hence M is a universal semi-POVM. □

For this universal semi-POVM $m(s)I$, we have $[m(s)I, m(t)I] = 0$ for all s and $t \in \Sigma^*$. However the following theorem guarantees an existence of a 'non-trivial' universal semi-POVM.

Theorem 3.6. *There exists a universal semi-POVM M such that $[M(s), M(t)] \neq 0$ for any distinct s and $t \in \Sigma^*$.*

Proof. We choose any one universal probability m, and choose any one pair of G and $H \in \mathrm{Her}_{\mathbb{Q}}(N)$ such that $0 < G, H \leqslant I$ and $[G, H] \neq 0$. We define $M\colon \Sigma^* \to \mathrm{Her}(N)$ by

$$M(s) = \frac{m(s)}{2}\left(2^{-\varphi(s)}G + H\right).$$

Then we see that, for any distinct s and $t \in \Sigma^*$,

$$[M(s), M(t)] = \frac{1}{4}m(s)m(t)\left(2^{-\varphi(s)} - 2^{-\varphi(t)}\right)[G, H] \neq 0.$$

Since m is a lower-computable semi-measure, M is shown to be a lower-computable semi-POVM. It follows from $0 < H$ that there is $c > 0$ such that $cI \leqslant H$. Thus $\frac{c}{2}m(s)I \leqslant M(s)$. Since $m(s)I$ is a universal semi-POVM, M is also a universal semi-POVM. □

The following theorem is more general form of our main result.

Theorem 3.7. *Let m be a universal probability, and let R be a lower-computable semi-POVM. Then the following (i) and (ii) hold:*

(i) There exists $c > 0$ such that, for any normalized $|\psi\rangle \in \mathbb{C}^N$ and any $s \in \Sigma^*$,

$$\langle\psi|R(s)|\psi\rangle \leq c\, m(s).$$

(ii) There exists $c > 0$ such that, for any density matrix $\rho \in \mathrm{Her}(N)$ and any $s \in \Sigma^*$,

$$\mathrm{tr}(\rho R(s)) \leq c\, m(s).$$

Proof. It follows from Theorem 3.5 that (i) holds. Using (i) and the spectral decomposition of ρ, we have (ii). □

In order to make more clear the physical implication of Theorem 3.7, we restrict our attention to a POVM on Σ^* which is computable. Informally, a POVM on Σ^* is computable if and only if one can compute all its POVM elements to any desired degree of precision. Thus the computability of a POVM is thought to be inherent in the case where one wants to perform a well-controlled quantum measurement described by the POVM. Using the following lemma, we have our main result about a computable POVM.

Lemma 3.8. *Let R be a semi-POVM on Σ^*. If R is computable then R is a lower-computable semi-POVM.*

Proof. Since R is computable, there exists a total recursive function $G\colon \Sigma^* \times \mathbb{N} \to M_N(\mathbb{C}_Q)$ such that, for all $s \in \Sigma^*$ and all $k \in \mathbb{N}$, $\|R(s) - G(s,k)\| < 2^{-k}$. We define $H\colon \Sigma^* \times \mathbb{N} \to M_N(\mathbb{C}_Q)$ by $H(s,k) = \frac{1}{2}\left\{G(s,k) + G(s,k)^\dagger\right\}$. Then H is a total recursive function and, for every $s \in \Sigma^*$ and every $k \in \mathbb{N}$, $H(s,k) \in \mathrm{Her}_{\mathbb{Q}}(N)$ and $\|R(s) - H(s,k)\| < 2^{-k}$. Thus we have $H(s,k) - 2^{-k}I \leqslant R(s)$ and $\lim_{k\to\infty} H(s,k) - 2^{-k}I = R(s)$. Hence the result follows. □

Theorem 3.9 (Main result). *Let R be a computable POVM on Σ^*. Then the following hold:*

(i) There exists $d \in \mathbb{N}$ such that, for any density matrix $\rho \in \mathrm{Her}(N)$ and any $s \in \Sigma^*$,

$$K(s) - d \leq -\log_2 \mathrm{tr}(\rho R(s)). \tag{4}$$

(ii) There exists $c > 0$ such that, for any density matrix $\rho \in \mathrm{Her}(N)$ and any $s \in \Sigma^*$,

$$\mathrm{tr}(\rho R(s)) \leq c\, P(s). \tag{5}$$

Proof. Theorem 3.9 immediately follows from Theorem 2.3, (ii) in Theorem 3.7, and Lemma 3.8. □

4 Optimality of universal semi-POVM

In this section we consider an optimality of a universal semi-POVM. By Theorem 3.5 we have the following theorem.

Theorem 4.1. *Let M be a universal semi-POVM. and let m be a universal probability. Then $M(s) \sim m(s)I$.*

The following theorem immediately follows from Theorem 4.1. This theorem is the most general form which represents the optimality of a universal-semi POVM from the view point of the probability of each measurement outcome.

Theorem 4.2. *Let m be a universal probability, and let M be a universal semi-POVM. Then there exist $c_1 > 0$ and $c_2 > 0$ such that, for any density matrix $\rho \in \mathrm{Her}(N)$ and any $s \in \Sigma^*$,*

$$c_1 m(s) \leq \mathrm{tr}(\rho M(s)) \leq c_2 m(s).$$

By Theorem 2.3 and Theorem 4.2, we have Theorem 4.3.

Theorem 4.3. *Let M be a universal semi-POVM. Then, for any density matrix $\rho \in \mathrm{Her}(N)$ and any $s \in \Sigma^*$,*

$$\begin{aligned} K(s) &= -\log_2 \mathrm{tr}(\rho M(s)) + O(1), \\ P(s) &\sim \mathrm{tr}(\rho M(s)). \end{aligned}$$

Thus, if we can perform the POVM measurement described by a universal semi-POVM, then we can achieve the upper bound $P(s)$ (or $2^{-K(s)}$) in Theorem 3.9 up to a multiplicative constant. However any universal semi-POVM is not computable (see Subsection 5.2). Moreover we can show that there is no computable semi-POVM on Σ^* which can achieve the upper bound $P(s)$ (or $2^{-K(s)}$) up to a multiplicative constant. Instead, by the definition of universal semi-POVM, we have the following theorem, which states that we can approximate any universal semi-POVM by a recursive sequence of semi-POVMs on Σ^* from below.

Theorem 4.4. *For any universal semi-POVM M, there exists a sequence $F_0, F_1, F_2, \ldots$ of semi-POVMs on Σ^* such that*

(i) $F_n(s) \in \mathrm{Her}_{\mathbb{Q}}(N)$ and $0 < F_n(s) \leqslant F_{n+1}(s) \leqslant M(s)$ for all $(n, s) \in \mathbb{N} \times \Sigma^*$,

(ii) the sequence $F_0, F_1, F_2, \ldots$ of functions uniformly converges to M, and

(iii) the mapping $\mathbb{N} \times \Sigma^* \ni (n, s) \longmapsto F_n(s)$ is a total recursive function.

Proof. Since M is a universal semi-POVM, by Theorem B.1 in Appendix B, there exists a total recursive function $g\colon \mathbb{N} \times \Sigma^* \to \mathrm{Her}_{\mathbb{Q}}(N)$ such that, for each $s \in \Sigma^*$, $\lim_{n\to\infty} g(n, s) = M(s)$ and $\forall\, n \in \mathbb{N}\ \ g(n, s) \leqslant g(n+1, s) \leqslant M(s)$. Note that $0 < M(s)$ for any $s \in \Sigma^*$. Thus, there exists a total recursive function $\tau\colon \mathbb{N} \times \Sigma^* \to \mathbb{N}$ such that, for each s and n, $\tau(n, s) < \tau(n+1, s)$ and $0 < g(\tau(n, s), s)$. We define the sequence $F_0, F_1, F_2, \ldots$ of semi-POVMs on Σ^* by $F_n(s) = g(\tau(n, s), s)$. It is then obvious that (i) and (iii) in Theorem 4.4 hold for this sequence. For any $\varepsilon > 0$, there is $s_0 \in \Sigma^*$ such that $\sum_{s > s_0} M(s) < \varepsilon I$, so we see that $\|F_n(s) - M(s)\| < \varepsilon$ for all $n \in \mathbb{N}$ and all $s > s_0$. On the other hand, it is easy to see that there is $n_0 \in \mathbb{N}$ such that, for all $n > n_0$ and all $s \leq s_0$, $\|F_n(s) - M(s)\| < \varepsilon$. Thus (ii) in Theorem 4.4 holds for the sequence $F_0, F_1, F_2, \ldots$ of functions. □

For the recursive sequence $F_0, F_1, F_2, \ldots$ of semi-POVMs on Σ^* given in Theorem 4.4, F_n is a computable semi-POVM on Σ^* for each $n \in \mathbb{N}$. However, since any universal semi-POVM is not a POVM on Σ^* (see Subsection 5.2) and $F_n(s) \leqslant M(s)$, F_n is not a POVM on Σ^* for each n. Instead, we can also consider the recursive sequence $G_1, G_2, G_3, \ldots$ of POVMs defined as follows: Each POVM element of G_n is labeled by a finite binary string less than or equal to $\varphi(n)$. For any $s < \varphi(n)$, $G_n(s)$ is defined as $F_n(s)$, and $G_n(\varphi(n))$ is defined as $I - \sum_{s<\varphi(n)} F_n(s)$. Then, since $F_n(s) \in \mathrm{Her}_{\mathbb{Q}}(N)$, any given $n \in \mathbb{N}^+$, one can calculate all POVM elements of G_n. Note that the POVM measurement described by G_n gives one of $n+1$ measurement outcomes, each of which is represented by a finite binary string less than or equal to $\varphi(n)$. By Theorem 4.4, we have the following:

- Any given $\varepsilon > 0$, for all sufficiently large $n \in \mathbb{N}^+$, if $s < \varphi(n)$ and ρ is a density matrix then $0 \leq \mathrm{tr}(\rho M(s)) - \mathrm{tr}(\rho G_n(s)) < \varepsilon$.

Thus, in the sense that the above statement holds, the recursive sequence $G_1, G_2, G_3, \ldots, G_n, \ldots$ of POVMs converges to the universal semi-POVM M from below as $n \to \infty$.

5 Other properties of universal semi-POVM

In this section we study the properties of universal semi-POVM further.

5.1 Matrix-valued algorithmic information theory

Let M be any one universal semi-POVM. The equation (3) suggests defining a matrix-valued Kolmogorov complexity $\mathcal{K}(s)$ of $s \in \Sigma^*$ by

$$\mathcal{K}(s) \equiv -\log_2 M(s). \tag{6}$$

For this definition of $\mathcal{K}$, it follows from Theorem 4.1 that

$$\mathcal{K}(s) = K(s)I + O(1). \tag{7}$$

Further we can define $\mathcal{K}(s,t)$, $\mathcal{K}(s|t)$, and $\mathcal{K}(s:t)$ in the same manner as the definitions of $K(s,t)$, $K(s|t)$, and $K(s:t)$, respectively. Then using (7) we see that all the relations in Theorem 2.5 hold for these $\mathcal{K}$'s in place of the K's.

Note that $K(s)$ is originally defined using the concept of program-size. Since $\mathcal{K}(s)$ is related to $K(s)$ through the equation (7), $\mathcal{K}(s)$ have the meaning of program-size in some weak sense. It is interesting if we can find a more concrete definition of $\mathcal{K}(s)$ using something like the concept of program-size instead of the equation (6). However, this is still open.

In order to measure the information content of a quantum state $|\psi\rangle \in \mathbb{C}^N$, [8] introduced the quantum algorithmic entropies $\underline{H}(|\psi\rangle)$ and $\overline{H}(|\psi\rangle)$ of $|\psi\rangle$ as $-\log_2\langle\psi|\boldsymbol{\mu}(N)|\psi\rangle$ and $-\langle\psi|(\log_2 \boldsymbol{\mu}(N))|\psi\rangle$, respectively, using his quantum universal semi-density matrix $\boldsymbol{\mu}$ (see Appendix A for its definition). In this behalf note that, for our universal semi-POVM M, the following holds for any normalized $|\psi\rangle \in \mathbb{C}^N$:

$$K(s) = -\log_2\langle\psi|M(s)|\psi\rangle + O(1) = -\langle\psi|(\log_2 M(s))|\psi\rangle + O(1).$$

Thus $-\log_2\langle\psi|M(s)|\psi\rangle$ and $-\langle\psi|(\log_2 M(s))|\psi\rangle$ are independent of $|\psi\rangle$ up to an additive constant. So it would seem difficult to measure the information content of a quantum state $|\psi\rangle$ using these quantities in the similar manner to [8], although such an attempt is not the purpose of this paper.

5.2 Relation to universal probability

We say $x \in \mathbb{C}^N$ is *computable* if each component of x is in the form of $a + ib$ where a and b are computable real numbers. Theorem 5.1 describes a property of a universal semi-POVM as a universal probability.

Theorem 5.1. *Let M be a universal semi-POVM, and let $x \in \mathbb{C}^N$ be computable with $\|x\| = 1$. Then the mapping $\Sigma^* \ni s \longmapsto x^\dagger M(s)x$ is a universal probability.*

Proof. Since x is computable, $x^\dagger M(s)x$ is shown to be a lower-computable semi-measure. Let m be a universal probability. Then, by Theorem 4.1, we see that $x^\dagger M(s)x \sim m(s)$. Thus the result follows. □

Let M be a universal semi-POVM. Then, by Theorem 5.1, each diagonal element $M_{ii}(s)$ of $M(s)$ is a universal probability as a function of s, and $\frac{1}{N}\operatorname{tr}(M(s))$ is also a universal probability as a function of s. Since any universal probability is not computable, any one diagonal element $M_{ii}(s)$ is not computable. Hence any universal semi-POVM is not computable. If M is a POVM on Σ^*, then, since M is a lower-computable semi-POVM, we can show that M is computable. Thus any universal semi-POVM is not a POVM on Σ^*.

5.3 Computable unitary invariance

We say $A \in M_N(\mathbb{C})$ is *computable* if each element of A is in the form of $a + ib$ where a and b are computable real numbers. The following theorem states an invariance of a POVM measurement described by a universal semi-POVM under computable unitary transformation on the quantum state being measured.

Theorem 5.2. *Let M be a universal semi-POVM, and let $U \in U(N)$ be computable. Then the mapping $\Sigma^* \ni s \longmapsto U^\dagger M(s)U$ is a universal semi-POVM.*

Proof. We note the property that $A \leqslant B \Longrightarrow X^\dagger AX \leqslant X^\dagger BX$ for any $A, B \in \mathrm{Her}(N)$ and any $X \in M_N(\mathbb{C})$. Since U is computable, $U^\dagger M(s)U$ is shown to be a lower-computable semi-POVM. Let m be a universal probability. Then, by Theorem 4.1, we see that $U^\dagger M(s)U \sim m(s)I \sim M(s)$. Thus the result follows. □

Let $U \in U(N)$ be a computable, and let $\mathcal{M}$ be a POVM measurement described by a universal semi-POVM. Suppose that, any given state ρ, we first evolve ρ by the unitary transformation U, and then perform the measurement $\mathcal{M}$ for the transformed state (i.e., $U\rho U^\dagger$). Then, by Theorem 5.2, the whole POVM measurement for ρ is shown to be still described by a universal semi-POVM.

Acknowledgments

The author is grateful to Hiroshi Imai and Keiji Matsumoto for their support.

References

[1] Bernstein E. and Vazirani U., Quantum complexity theory, *SIAM J. Comput.*, **26** (1997), 1411–1473.

[2] Berthiaume A., van Dam W., and Laplante S., Quantum Kolmogorov complexity, *J. Compute. System Sci.*, **63** (2001), 201–221.

[3] Bhatia R., *Matrix Analysis*, Springer, New York, 1996.

[4] Calude C. S., Hertling P. H., Khoussainov B., and Wang Y., Recursively enumerable reals and Chaitin Ω numbers, *Theoret. Comput. Sci.*, **255** (2001), pp.125–149.

[5] Calude C. S., *Information and Randomness: An Algorithmic Perspective*, 2nd Edition, Revised and Extended, Springer, Berlin, 2002.

[6] Chaitin G. J., A theory of program size formally identical to information theory, *J. Assoc. Comput. Mach.*, **22** (1975), pp.329–340.

[7] Chaitin G. J., Incompleteness theorems for random reals, *Adv. in Appl. Math.*, **8** (1987), pp.119–146.

[8] Gács P., Quantum algorithmic entropy, *J. Phys. A: Math. Gen.*, **34** (2001), pp.6859–6880.

[9] Horn R. A. and Johnson C. R., *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[10] Kučera A. and Slaman T. A., Randomness and recursive enumerability, *SIAM J. Comput.*, **31** (2001), pp.199–211.

[11] Nielsen M. A. and Chuang I. L., *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, 2000.

[12] Preskill J., *Quantum Computation*, 2000. Course notes available at URL: `http://www.theory.caltech.edu/people/preskill/ph229/`.

[13] Vitányi P. M. B., Quantum Kolmogorov complexity based on classical descriptions, *IEEE Trans. Inform. Theory*, **47** (2001), pp.2464–2479.

[14] Li M. and Vitányi P. M. B., *An Introduction to Kolmogorov Complexity and Its Applications*, Second Edition, Springer, New York, 1997.

A Quantum universal semi-density matrix

We reproduce the definition of quantum universal semi-density matrix from [8] as follows.

Definition A.1. *Let* $\sigma\colon \mathbb{N}^+ \to \bigcup_{N\geq 1} \mathrm{Her}(N)$. *We say* σ *is a lower semicomputable semi-density matrix if* σ *satisfies the following conditions:*

(i) For each $N \in \mathbb{N}^+$, $0 \leqslant \sigma(N) \in \mathrm{Her}(N)$ *and* $\mathrm{tr}(\sigma(N)) \leq 1$.

(ii) There exists a total recursive function $f\colon \mathbb{N}^+ \times \mathbb{N} \to \bigcup_{N\geq 1} \mathrm{Her}_{\mathbb{Q}}(N)$ *such that, for each* $N \in \mathbb{N}^+$, $\lim_{k\to\infty} f(N,k) = \sigma(N)$ *and* $\forall\, k \in \mathbb{N}\ \ f(N,k) \in \mathrm{Her}_{\mathbb{Q}}(N)\ \&\ f(N,k) \leqslant f(N,k+1)$.

Definition A.2. *Let* $\boldsymbol{\mu}$ *be a lower semicomputable semi-density matrix. We say* $\boldsymbol{\mu}$ *is a quantum universal semi-density matrix if for any lower semicomputable semi-density matrix* σ*, there exists a real number* $c > 0$ *such that, for all* $N \in \mathbb{N}^+$, $c\,\sigma(N) \leqslant \boldsymbol{\mu}(N)$.

Theorem A.3. *If* $\boldsymbol{\mu}$ *is a quantum universal semi-density matrix, then* $\mathrm{tr}(\boldsymbol{\mu}(N)) < 1$ *for all but finitely many* $N \in \mathbb{N}^+$.

Proof. Since $\boldsymbol{\mu}$ is a lower semicomputable semi-density matrix, there exists a total recursive function f on $\mathbb{N}^+ \times \mathbb{N}$ such that, for each $N \in \mathbb{N}^+$, $\lim_{k\to\infty} f(N,k) = \boldsymbol{\mu}(N)$ and $\forall\, k \in \mathbb{N}\ f(N,k) \in \mathrm{Her}_{\mathbb{Q}}(N)$ & $f(N,k) \leqslant \boldsymbol{\mu}(N)$. Let $\boldsymbol{\mu}_{ii}(N)$ be the (i,i)-element of $\boldsymbol{\mu}(N)$, and let $f_{ii}(N,k)$ be the (i,i)-element of $f(N,k)$. Then, since $\mathrm{tr}(\boldsymbol{\mu}(N)) \leq 1$, we see that $\boldsymbol{\mu}_{ii}(N) \leq 1 - \sum_{j\neq i} f_{jj}(N,k)$. Especially, for any N with $\mathrm{tr}(\boldsymbol{\mu}(N)) = 1$, we have $\boldsymbol{\mu}_{ii}(N) = \lim_{k\to\infty} 1 - \sum_{j\neq i} f_{jj}(N,k)$. On the other hand, it follows from $\sum_{i=1}^{N} \boldsymbol{\mu}_{ii}(N) \leq 1$ that $\min\{\boldsymbol{\mu}_{ii}(N) \mid 1 \leq i \leq N\} \leq 1/N$. Therefore, any given $\varepsilon > 0$, for each sufficiently large N, there is i such that $1 \leq i \leq N$ and $\boldsymbol{\mu}_{ii}(N) < \varepsilon$.

Now, contrary to Theorem A.3, let us assume that, for infinitely many $N \in \mathbb{N}^+$, $\mathrm{tr}(\boldsymbol{\mu}(N)) = 1$. Then, given $\varepsilon \in \mathbb{Q}^+$, by checking for each (N,i,k) in an exhaustive order whether $1 - \sum_{j\neq i} f_{jj}(N,k) < \varepsilon$ holds or not, one can find (N,i) such that $\boldsymbol{\mu}_{ii}(N) < \varepsilon$. Let m be any one universal probability, and we define the lower semicomputable semi-density matrix σ by $\sigma(N) = \mathrm{diag}(m(\varphi^{-1}(1)), \ldots, m(\varphi^{-1}(N)))$. Then, since $\boldsymbol{\mu}$ is a quantum universal semi-density matrix, for this σ, there is $c_\sigma > 0$ such that if $1 \leq i \leq N$ then $c_\sigma m(\varphi^{-1}(i)) \leq \boldsymbol{\mu}_{ii}(N)$. It follows that there exists a total recursive function $\tau\colon \mathbb{N} \to \Sigma^*$ such that, for any $n \in \mathbb{N}$, $m(\tau(n)) \leq 2^{-n}$. This contradicts Theorem 2.4. Thus we have Theorem A.3. □

B On the definition of lower-computable semi-POVM

The following theorem guarantees that one can equivalently assume that $f(n,s)$ converging to $R(s)$ is non-decreasing in Definition 3.2.

Theorem B.1. *R is a lower-computable semi-POVM if and only if R is a semi-POVM on Σ^* and there exists a total recursive function $f\colon \mathbb{N} \times \Sigma^* \to \mathrm{Her}_{\mathbb{Q}}(N)$ such that for each $s \in \Sigma^*$, $\lim_{n\to\infty} f(n,s) = R(s)$ and $\forall\, n \in \mathbb{N}\ \ f(n,s) \leqslant f(n+1,s)$.*

For the proof of Theorem B.1 we need the following lemma, which is an elementary result of linear algebra.

Lemma B.2. *For any $A \in \mathrm{Her}(N)$, $0 \leqslant A$ if and only if all principal minors of A are non-negative.*

By Lemma B.2, given A and B in $\mathrm{Her}_{\mathbb{Q}}(N)$, one can effectively check whether $A \leqslant B$ holds or not.

Proof of Theorem B.1. Assume that R is a semi-POVM on Σ^* and there exists a total recursive function $f\colon \mathbb{N} \times \Sigma^* \to \mathrm{Her}_{\mathbb{Q}}(N)$ such that for each $s \in \Sigma^*$, $\lim_{n\to\infty} f(n,s) = R(s)$ and $\forall\, n \in \mathbb{N}\ \ f(n,s) \leqslant R(s)$. Let $g\colon \mathbb{N} \times \Sigma^* \to \mathrm{Her}_{\mathbb{Q}}(N)$ be a total recursive function such that $g(n,s) = f(n,s) - 2^{-n}I$. Then, for each $s \in \Sigma^*$, $\lim_{n\to\infty} g(n,s) = R(s)$ and $\forall\, n \in \mathbb{N}\ g(n,s) < R(s)$. Thus, for each s and n, there is a positive real number c such that $cI \leqslant R(s) - g(n,s)$, and then, for this c, there is an $m \in \mathbb{N}$ such that $m > n$ and $R(s) - g(m,s) \leqslant cI$. So we have $g(n,s) \leqslant g(m,s)$. Thus, given s and n, by checking $g(n,s) \leqslant g(k,s)$ for each $k > n$ in increasing order, one can finally find an m with $g(n,s) \leqslant g(m,s)$. Therefore there exists a total recursive function $\tau\colon \mathbb{N}\times\Sigma^* \to \mathbb{N}$ such that, for each s and n, $\tau(n,s) < \tau(n+1,s)$ and $g(\tau(n,s),s) \leqslant g(\tau(n+1,s),s)$. We define a total recursive function $h\colon \mathbb{N} \times \Sigma^* \to \mathrm{Her}_{\mathbb{Q}}(N)$ by $h(n,s) = g(\tau(n,s),s)$. Then, for each $s \in \Sigma^*$, $\lim_{n\to\infty} h(n,s) = R(s)$ and $\forall\, n \in \mathbb{N}\ \ h(n,s) \leqslant h(n+1,s)$. Hence, R is a lower-computable semi-POVM.

The other implication is obvious. Thus the theorem is obtained. □